codecademy

Capstone: Warby Parker

Learn SQL from Scratch

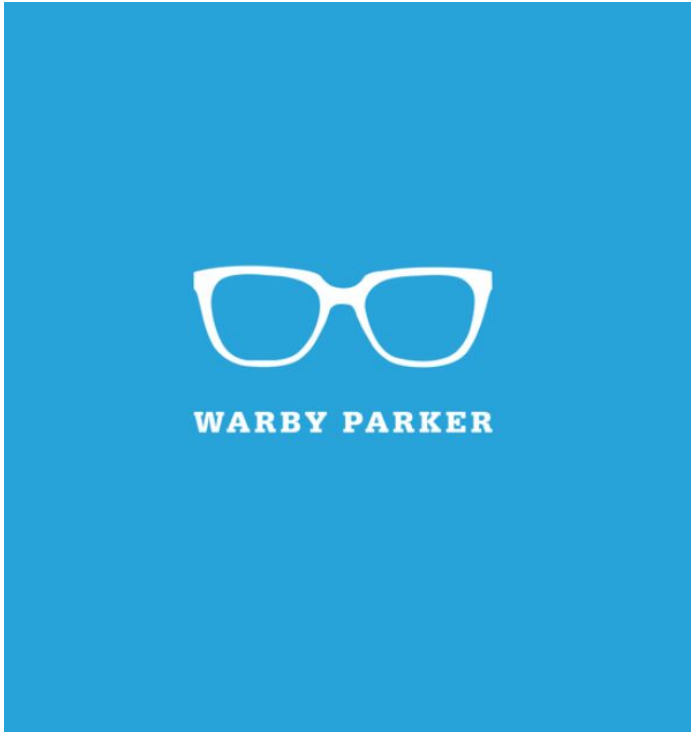
Michael Covey

September 2018

OVERVIEW

Warby Parker offers shoppers the opportunity to learn about their eyeglass preferences through an **online quiz.**

After taking the quiz, WP sends shoppers a set of **tester eyeglasses to try on at home for free**. Shoppers can then decide for themselves which pair to purchase.



SHOPPER FLOW

1. Online survey

Shoppers take a short quiz on their preferred eyeglass styles, their gender, and the last time they took an eye exam.

2. At-home trial

Based on the survey, shoppers are sent either 3 or 5 pairs of tester eyeglasses (internal A/B test) to try on for free.

3. Purchase

After the trial is over, shoppers can elect to purchase a pair of eyeglasses from Warby Parker's website.

OBJECTIVE

Help Warby Parker answer the following questions:

1. Which questions on the online quiz are most **prone to attrition**?
2. What are the **conversion rates** between the online quiz, trying on at home, and purchase?
3. Is purchase likelihood (and price paid) affected by **how many pairs** are sent during the home trial period?
4. **Gender** deep-dives
    a. Do men and women **prefer different styles**?
    b. Do men and women have different **conversion rates**?
5. Are conversion rates affected by the **styles selected during the online quiz**?

```sql
SELECT *
FROM survey
LIMIT 10;
```

SETUP

Selecting all columns from the first 10 rows of the `survey` table.

question	user_id	response
1. What are you looking for?	005e7f99-d48c-4fce-b605-10506c85aaf7	Women's Styles
2. What's your fit?	005e7f99-d48c-4fce-b605-10506c85aaf7	Medium
3. Which shapes do you like?	00a556ed-f13e-4c67-8704-27e3573684cd	Round
4. Which colors do you like?	00a556ed-f13e-4c67-8704-27e3573684cd	Two-Tone
1. What are you looking for?	00a556ed-f13e-4c67-8704-27e3573684cd	I'm not sure. Let's skip it.
2. What's your fit?	00a556ed-f13e-4c67-8704-27e3573684cd	Narrow
5. When was your last eye exam?	00a556ed-f13e-4c67-8704-27e3573684cd	<1 Year
3. Which shapes do you like?	00bf9d63-0999-43a3-9e5b-9c372e6890d2	Square
5. When was your last eye exam?	00bf9d63-0999-43a3-9e5b-9c372e6890d2	<1 Year
2. What's your fit?	00bf9d63-0999-43a3-9e5b-9c372e6890d2	Medium

SURVEY FUNNEL

```sql
SELECT question AS 'Question',
  COUNT(user_id) AS 'Responses'
FROM survey
GROUP BY 1;
```

This code builds a funnel for the survey, which can help us understand the specific questions at which users stop taking the survey.

Below, you can see that 500 users started taking the survey, but **only 270 finished it (54%)**.

Question	Responses
1. What are you looking for?	500
2. What's your fit?	475
3. Which shapes do you like?	380
4. Which colors do you like?	361
5. When was your last eye exam?	270

SURVEY FUNNEL

The last question is causing the highest levels of survey attrition (“drop-off”). Users may not remember when they last had an eye exam. Warby Parker should consider **providing eye exam resources** to encourage users to stay on the platform.

Similarly, the shapes question lost 20% of respondents. Warby Parker should consider providing resources to **help users understand which shapes will look best on their face**.

Question	Response Count	% completing
1. What are you looking for?	500	-
2. What's your fit?	475	95%
3. Which shapes do you like?	380	***80%***
4. Which colors do you like?	361	95%
5. When was your last eye exam?	270	***75%***

SETUP

```sql
SELECT *
FROM quiz
LIMIT 5;

SELECT *
FROM home_try_on
LIMIT 5;

SELECT *
FROM purchase
LIMIT 5;
```

Viewing first five rows of the `quiz`, `home_try_on`, and `purchase` tables.

user_id	style	fit	shape	color
4e8118dc-bb3d-49bf-85fc-cca8d83232ac	Women's Styles	Medium	Rectangular	Tortoise
291f1cca-e507-48be-b063-002b14906468	Women's Styles	Narrow	Round	Black
75122300-0736-4087-b6d8-c0c5373a1a04	Women's Styles	Wide	Rectangular	Two-Tone
75bc6ebd-40cd-4e1d-a301-27ddd93b12e2	Women's Styles	Narrow	Square	Two-Tone
ce965c4d-7a2b-4db6-9847-601747fa7812	Women's Styles	Wide	Rectangular	Black

user_id	number_of_pairs	address
d8addd87-3217-4429-9a01-d56d68111da7	5 pairs	145 New York 9a
f52b07c8-abe4-4f4a-9d39-ba9fc9a184cc	5 pairs	383 Madison Ave
8ba0d2d5-1a31-403e-9fa5-79540f8477f9	5 pairs	287 Pell St
4e71850e-8bbf-4e6b-accc-49a7bb46c586	3 pairs	347 Madison Square N
3bc8f97f-2336-4dab-bd86-e391609dab97	5 pairs	182 Cornelia St

user_id	product_id	style	model_name	color	price
00a9dd17-36c8-430c-9d76-df49d4197dcf	8	Women's Styles	Lucy	Jet Black	150
00e15fe0-c86f-4818-9c63-3422211baa97	7	Women's Styles	Lucy	Elderflower Crystal	150
017506f7-aba1-4b9d-8b7b-f4426e71b8ca	4	Men's Styles	Dawes	Jet Black	150
0176bfb3-9c51-4b1c-b593-87edab3c54cb	10	Women's Styles	Eugene Narrow	Rosewood Tortoise	95
01fdf106-f73c-4d3f-a036-2f3e2ab1ce06	8	Women's Styles	Lucy	Jet Black	150

```sql
SELECT
DISTINCT q.user_id,
h.user_id IS NOT NULL AS 'is_home_try_on',
h.number_of_pairs AS 'number_of_pairs',
p.user_id IS NOT NULL AS 'is_purchase'
FROM quiz AS 'q'
LEFT JOIN home_try_on AS 'h'
  ON h.user_id = q.user_id
LEFT JOIN purchase AS 'p'
  ON p.user_id = h.user_id
LIMIT 10;
```

MORE SETUP

Joining the `quiz`, `home_try_on`, and `purchase` tables in order to make conversion analysis easier.

user_id	is_home_try_on	number_of_pairs	is_purchase
4e8118dc-bb3d-49bf-85fc-cca8d83232ac	1	3 pairs	0
291f1cca-e507-48be-b063-002b14906468	1	3 pairs	1
75122300-0736-4087-b6d8-c0c5373a1a04	0	Ø	0
75bc6ebd-40cd-4e1d-a301-27ddd93b12e2	1	5 pairs	0
ce965c4d-7a2b-4db6-9847-601747fa7812	1	3 pairs	1
28867d12-27a6-4e6a-a5fb-8bb5440117ae	1	5 pairs	1
5a7a7e13-fbcf-46e4-9093-79799649d6c5	0	Ø	0
0143cb8b-bb81-4916-9750-ce956c9f9bd9	0	Ø	0
a4ccc1b3-cbb6-449c-b7a5-03af42c97433	1	5 pairs	0
b1dded76-cd60-4222-82cb-f6d464104298	1	3 pairs	0

```sql
WITH funnel AS(
SELECT
 DISTINCT q.user_id,
 h.user_id IS NOT NULL AS 'is_home_try_on',
 h.number_of_pairs AS 'number_of_pairs',
 p.user_id IS NOT NULL AS 'is_purchase'
FROM quiz AS 'q'
LEFT JOIN home_try_on AS 'h'
  ON h.user_id = q.user_id
LEFT JOIN purchase AS 'p'
  ON p.user_id = h.user_id)
SELECT COUNT(*) AS 'All users',
SUM(is_home_try_on) AS 'Home Try-Ons',
SUM(is_purchase) AS 'Purchasers'
FROM funnel;
```

All users	Home Try-Ons	Purchasers
1000	750	495

CONVERSION

- 1000 shoppers took the quiz.
- 750 shoppers requested a home try-on (75%).
- 495 purchased a pair of glasses (~50%).

A/B TEST

Warby Parker wanted to see if sending shoppers either 3 or 5 pairs of glasses during the try-on period would affect purchase likelihood.

This is the code to answer this question. (The output is useless without additional context.)

```sql
--Counts purchasers who got either 3 or 5 home try-on pairs
WITH funnel AS(
SELECT
 DISTINCT q.user_id,
 h.user_id IS NOT NULL AS 'is_home_try_on',
 h.number_of_pairs AS 'number_of_pairs',
 p.user_id IS NOT NULL AS 'is_purchase'
FROM quiz AS 'q'
LEFT JOIN home_try_on AS 'h'
  ON h.user_id = q.user_id
LEFT JOIN purchase AS 'p'
  ON p.user_id = h.user_id)
SELECT number_of_pairs, SUM(is_purchase) AS 'Count'
FROM funnel
GROUP BY number_of_pairs
HAVING number_of_pairs IS NOT NULL;
```

```sql
--Count total number of home try-ons (regardless of purchase)
SELECT number_of_pairs, COUNT(*) AS 'Count'
FROM home_try_on
GROUP BY 1;
```

A/B TEST

Dividing the number of purchasers by the number of try-ons, we can calculate the conversion rate.

Here, we see strong evidence that giving shoppers more pairs of glasses during the trial period results in a greater likelihood of purchase.

Pair Count	Total Try-ons	Total purchasers	Conversion
3 pairs	379	201	53%
5 pairs	371	294	79%

```sql
--Counts purchasers who got either 3 or 5 home try-on pairs
WITH funnel AS(
SELECT
 DISTINCT q.user_id, p.price,
 h.user_id IS NOT NULL AS 'is_home_try_on',
 h.number_of_pairs AS 'number_of_pairs',
 p.user_id IS NOT NULL AS 'is_purchase'
FROM quiz AS 'q'
LEFT JOIN home_try_on AS 'h'
  ON h.user_id = q.user_id
LEFT JOIN purchase AS 'p'
  ON p.user_id = h.user_id)
SELECT number_of_pairs, AVG(price)
FROM funnel
GROUP BY number_of_pairs
HAVING number_of_pairs IS NOT NULL;
```

number_of_pairs	AVG(price)
3 pairs	113.258706467662
5 pairs	112.34693877551

A/B TEST

However, the number of pairs during the trial period does not seem to affect how much shoppers spend on glasses.

Regardless of how many trial glasses sent, shoppers spend about $112 on a pair of glasses (if they decide to buy).

```
WITH funnel AS(
SELECT
 DISTINCT q.style AS 'gender', q.user_id,
 h.user_id IS NOT NULL AS 'is_home_try_on',
 h.number_of_pairs AS 'number_of_pairs',
 p.user_id IS NOT NULL AS 'is_purchase'
FROM quiz AS 'q'
LEFT JOIN home_try_on AS 'h'
  ON h.user_id = q.user_id
LEFT JOIN purchase AS 'p'
  ON p.user_id = h.user_id)
SELECT gender, SUM(is_purchase)
FROM funnel
GROUP BY gender;
```

GENDER

Next, we explored how the gender of the shopper might influence style preferences, as well as conversion rate.

(Analysis on following pages)

```
--Quiz-takers broken out by gender
SELECT style, COUNT(*)
FROM quiz
GROUP BY 1;
```

gender	SUM(is_purchase)
I'm not sure. Let's skip it.	0
Men's Styles	243
Women's Styles	252

style	COUNT(*)
I'm not sure. Let's skip it.	99
Men's Styles	432
Women's Styles	469

GENDER - CONVERSION RATES

Both men and women were about equally likely to purchase a pair of glasses.

However, for those who skipped the gender style question in the survey, 0% went on to purchase.

Gender	Conversion
Men	56%
Women	54%
Unspecified	***0%***

*WP may be overlooking an opportunity to **provide gender-neutral offerings** to shoppers who have a difficult time answering the gender survey question.*

GENDER - STYLE PREFERENCES

```sql
--Gender X Fit crosstab
SELECT style, fit, COUNT(fit)
FROM quiz
WHERE NOT style LIKE '%skip%'
GROUP BY 1, 2;
```

Fit	Men	Women
Narrow	40%	40%
Medium	33%	28%
Wide	18%	22%
Not sure	9%	10%

```sql
--Gender X Shape crosstab
SELECT style, shape, COUNT(fit)
FROM quiz
WHERE NOT style LIKE '%skip%'
GROUP BY 1, 2;
```

Shape	Men	Women
Rectangular	41%	39%
Square	31%	34%
Round	19%	17%
No preference	10%	10%

```sql
--Gender X Color crosstab
SELECT style, color, COUNT(*)
FROM quiz
WHERE NOT style LIKE '%skip%'
GROUP BY 1, 2;
```

Color	Men	Women
Tortoise	30%	30%
Black	28%	27%
Crystal	19%	23%
Two-tone	13%	8%
Neutral	10%	12%

Men and women tended to have **similar fit, shape, and color preferences**.

We’ve established that men and women have similar style/fit/color preferences.

Do initial style/fit/color preferences, as indicated in the survey, *affect purchase likelihood?*

The following three slides explore this possibility.

```sql
--Purchasers based on fit selection of quiz-takers
WITH funnel AS(
SELECT
 DISTINCT q.fit AS 'Fit', q.user_id,
 h.user_id IS NOT NULL AS 'is_home_try_on',
 h.number_of_pairs AS 'number_of_pairs',
 p.user_id IS NOT NULL AS 'is_purchase'
FROM quiz AS 'q'
LEFT JOIN home_try_on AS 'h'
  ON h.user_id = q.user_id
LEFT JOIN purchase AS 'p'
  ON p.user_id = h.user_id)
SELECT Fit, SUM(is_purchase)
FROM funnel
GROUP BY Fit;
--All fit selections regardless of purchase
SELECT fit, COUNT(*)
FROM quiz
GROUP BY 1;
```

Fit	Conversion
Narrow	47%
Medium	50%
Wide	53%
Not sure	51%

STYLE: *FIT* CONVERSION RATES

Shoppers who selected Wide as their fit preference in the survey were slightly more likely to go on to make a purchase.

```sql
--Purchasers based on shape selection of quiz-takers
WITH funnel AS(
SELECT
 DISTINCT q.shape AS 'Shape', q.user_id,
 h.user_id IS NOT NULL AS 'is_home_try_on',
 h.number_of_pairs AS 'number_of_pairs',
 p.user_id IS NOT NULL AS 'is_purchase'
FROM quiz AS 'q'
LEFT JOIN home_try_on AS 'h'
  ON h.user_id = q.user_id
LEFT JOIN purchase AS 'p'
  ON p.user_id = h.user_id)
SELECT Shape, SUM(is_purchase)
FROM funnel
GROUP BY Shape;
--All shape selections regardless of purchase
SELECT shape, COUNT(*)
FROM quiz
GROUP BY 1;
```

Shape	Conversion
Rectangular	48%
Round	53%
Square	48%
No preference	55%

STYLE: *SHAPE* CONVERSION RATES

Shoppers who selected **Rectangular** or **Square** in the survey were less likely to go on to make a purchase.

```sql
--Purchasers based on color selection of quiz-takers
WITH funnel AS(
SELECT
 DISTINCT q.color AS 'Color', q.user_id,
 h.user_id IS NOT NULL AS 'is_home_try_on',
 h.number_of_pairs AS 'number_of_pairs',
 p.user_id IS NOT NULL AS 'is_purchase'
FROM quiz AS 'q'
LEFT JOIN home_try_on AS 'h'
  ON h.user_id = q.user_id
LEFT JOIN purchase AS 'p'
  ON p.user_id = h.user_id)
SELECT Color, SUM(is_purchase)
FROM funnel
GROUP BY Color;
--All color selections regardless of purchase
SELECT color, COUNT(*)
FROM quiz
GROUP BY 1;
```

Color	Conversion
Black	54%
Crystal	50%
Neutral	***42%***
Tortoise	49%
Two-Tone	47%

STYLE: *COLOR* CONVERSION RATES

Shoppers who selected **Neutral color** in the survey were much less likely to go on to make a purchase.

Thank you!

Questions?

michaelccovey@gmail.com